

University of Groningen

A New Condition for Transitivity of Probabilistic Support

Atkinson, David; Peijnenburg, Jeanne

Published in:
Erkenntnis

DOI:
[10.1007/s10670-020-00349-7](https://doi.org/10.1007/s10670-020-00349-7)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2023

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):
Atkinson, D., & Peijnenburg, J. (2023). A New Condition for Transitivity of Probabilistic Support. *Erkenntnis*, 88, 253-265. <https://doi.org/10.1007/s10670-020-00349-7>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.



A New Condition for Transitivity of Probabilistic Support

David Atkinson¹ · Jeanne Peijnenburg¹

Received: 19 January 2020 / Accepted: 14 December 2020 / Published online: 19 March 2021
© The Author(s) 2021

Abstract

As is well known, implication is transitive but probabilistic support is not. Eells and Sober, followed by Shogenji, showed that screening off is a sufficient constraint for the transitivity of probabilistic support. Moreover, this screening off condition can be weakened without sacrificing transitivity, as was demonstrated by Suppes and later by Roche. In this paper we introduce an even weaker sufficient condition for the transitivity of probabilistic support, in fact one that can be made as weak as one wishes. We explain that this condition has an interesting property: it shows that transitivity is retained even though the Simpson paradox reigns. We further show that by adding a certain restriction the condition can be turned into one that is both sufficient and necessary for transitivity.

Keywords Probabilistic support · Transitivity · Screening off · Sufficient condition · Necessary condition · Simpson paradox

1 Introduction

We say that proposition p probabilistically supports proposition q , and that q probabilistically supports r if and only if

$$P(q|p) - P(q) > 0 \quad \text{and} \quad P(r|q) - P(r) > 0. \quad (1)$$

Unlike implication or entailment, probabilistic support is not transitive; in general it does not follow from (1) that p also supports r , i.e. that

$$P(r|p) - P(r) > 0. \quad (2)$$

✉ Jeanne Peijnenburg
jeanne.peijnenburg@rug.nl

David Atkinson
d.atkinson@rug.nl

¹ Faculty of Philosophy, University of Groningen, Oude Boteringestraat 52, 9712 GL Groningen, Netherlands

In 2003 Tomoji Shogenji proved that screening off is sufficient for making probabilistic support transitive: if p supports q and q supports r and there is screening off, then p supports r .¹ In 2012 William Roche described a weak version of screening off, and he demonstrated that it still suffices for transitivity.² In this paper we weaken Roche's condition further; and we demonstrate that the transitivity of probabilistic support still holds.

Our argument is set up as follows. In Sect. 2 we recall the two versions of screening off: the normal version, which is a particular Markov condition, and the weak variant of Roche. An identity that Tomoji Shogenji developed in 2017 enables us to demonstrate in a particularly succinct manner that both versions suffice for transitivity. In Sect. 3 we show that there exists an even weaker condition, one that still guarantees the transitivity of probabilistic support. As we explain, our new condition covers a continuum of conditions, each of which is weaker than Roche's, and each of which preserves transitivity. Section 4 highlights an interesting property of the new condition: it guarantees transitivity of probabilistic support even if the Simpson paradox obtains. The Simpson effect, as we prefer to call this paradox, implies that p disconfirms r conditionally on q and also disconfirms r conditionally on $\neg q$, while p still confirms r unconditionally. Rather surprisingly, as we show, this effect can coexist with transitivity: p confirms q , q confirms r , and p confirms r . In Sect. 5 we illustrate this coexistence with a well-known medical example. We conclude in Sect. 6 by constraining our new sufficient condition in such a way that it is also necessary for the transitivity of probabilistic support.

2 Normal and Weak Screening Off

Our first task is to find a condition, C , that is sufficient for transitivity in the sense: "If C , then if (1) then (2)", or equivalently:

$$(1) \rightarrow \left\{ C \rightarrow (2) \right\}. \quad (3)$$

In 2003 Tomoji Shogenji showed that (3) is valid if for C we fill in the condition of screening off, which here means that q screens off p from r , that is,

$$P(r|q \wedge p) - P(r|q) = 0 \quad \text{and} \quad P(r|\neg q \wedge p) - P(r|\neg q) = 0. \quad (4)$$

Moreover, in 2012 William Roche demonstrated that (3) remains valid if we weaken the condition (4) to

$$P(r|q \wedge p) - P(r|q) \geq 0 \quad \text{and} \quad P(r|\neg q \wedge p) - P(r|\neg q) \geq 0, \quad (5)$$

¹ Shogenji (2003). A similar proof concerning probabilistic causality had been given earlier in Eells and Sober (1983).

² Roche (2012). Unbeknownst to Roche, a proof had already been provided in Suppes (1986).

where the equals signs have been replaced by inequalities.³ A good way to see that both (4) and (5) are sufficient conditions for the transitivity of probabilistic support is by means of an illuminating analysis that Shogenji gave of what is in our notation $P(r|p) - P(r)$.⁴ We will not reproduce here all the actual steps of Shogenji’s careful argument. For our purpose it is enough to say that he derives a relation which in our reworking is the following:

$$P(r|p) - P(r) = \mu(p, r; q) + \mu(p, r; \neg q) + \tau(p, r; q), \tag{6}$$

where

$$\begin{aligned} \mu(p, r; q) &= [P(r|q \wedge p) - P(r|q)]P(q|p) \\ \mu(p, r; \neg q) &= [P(r|\neg q \wedge p) - P(r|\neg q)]P(\neg q|p) \\ \tau(p, r; q) &= \frac{P(q|p) - P(q)}{1 - P(q)} (P(r|q) - P(r)). \end{aligned}$$

Equation (6) is an identity: it is valid for any propositions p, q and r , irrespective of whether there is screening off, or even probabilistic support.⁵ The expressions for μ and τ contain four Carnap measures of confirmation. For example, $P(r|q) - P(r)$ is Carnap’s difference measure.⁶ The square brackets in the definitions of μ are the Carnap measures of confirmation, conditional on q and $\neg q$ respectively. The additional factors $P(q|p)$, $P(\neg q|p)$, and the division by $1 - P(q)$, are inert, in the sense that none of them is zero (see footnote 3).

If (4) holds, then $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ are both zero:

$$\mu(p, r; q) = \mu(p, r; \neg q) = 0. \tag{7}$$

Since neither $P(q|p)$ nor $P(\neg q|p)$ is zero, (7) is equivalent to (4). We will call (7) the condition of *normal* screening off. Under this condition, (6) reduces to:

$$P(r|p) - P(r) = \tau(p, r; q). \tag{8}$$

³ It is assumed that all the conditional probabilities are well defined, which implies that $P(q)$, $P(\neg q)$, $P(q \wedge p)$, and $P(\neg q \wedge p)$ are all non-zero.

⁴ Shogenji (2017).

⁵ The actual identity that Shogenji proves in Appendix A of his 2017 paper is the following:

$$\begin{aligned} P(z|x) - P(z) &= [P(z|y) - P(z)][P(y|x) - P(y)] + [P(z|\neg y) - P(z)][P(\neg y|x) - P(\neg y)] \\ &\quad + P(y|x)[P(z|y \wedge x) - P(z|y)] + P(\neg y|x)[P(z|\neg y \wedge x) - P(z|\neg y)]. \end{aligned}$$

The third term on the right here corresponds to our $\mu(p, r; q)$ and the fourth term to $\mu(p, r; \neg q)$, with the identification of his x, y, z with our p, q, r , respectively. After an elementary transformation the sum of the first two terms becomes our $\tau(p, r; q)$.

⁶ This is Carnap’s measure D , in his words the degree to which r is made firmer by q , see Carnap (1962, p. xvi). The factor outside the large parentheses in $\tau(p, r; q)$ is actually the z -measure in Crupi et al. (2007) in the case $P(q|p) > P(q)$.

If p supports q and q supports r , then the right-hand side of (8) is positive. Therefore the left-hand side is positive: p supports r . So with the help of (6) we see that normal screening off suffices for transitivity.

But (6) also allows us to see that there is transitivity under (5) as well, for the inequalities (5) imply that both $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ are non-negative:

$$\mu(p, r; q) \geq 0 \quad \text{and} \quad \mu(p, r; \neg q) \geq 0. \quad (9)$$

If p supports q and q supports r , then $\tau(p, r; q)$ is strictly positive, as we have just seen, and so with (9) the right-hand side of (6) is still positive. In fact, we do not need $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ to be separately non-negative, it is clear from (6) that it is enough if their sum is non-negative:

$$\mu(p, r; q) + \mu(p, r; \neg q) \geq 0. \quad (10)$$

Condition (10) is a little weaker than (9) because it allows for the possibility that one of $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ may be negative, on condition that the other is positive and sufficiently large to guarantee that the sum is equal to or greater than zero. We will call condition (10) *weak screening off* or *w-screening off* for short. If (1), then with w-screening off it follows that p supports r : $P(r|p) > P(r)$.

Conclusion: C in (3) may be identified with either (7), the normal screening off condition, or with (10), the condition of w-screening off. In both cases, transitivity has been ensured.

3 Very Weak Screening Off

In this section we will weaken w-screening off still further while retaining the transitivity. Our new condition, which we may call *very weak screening off* or *vw-screening off* for short, has to satisfy two requirements:

- (i) it is weaker than w-screening off
- (ii) it makes the left-hand side of (6) positive, so that p supports r , and transitivity has been achieved.

Clearly requirement (i) is fulfilled if $\mu(p, r; q) + \mu(p, r; \neg q) < 0$, for this is a situation explicitly ruled out by (10), the condition of w-screening off. As to requirement (ii), this can still be fulfilled even if the sum $\mu(p, r; q) + \mu(p, r; \neg q)$ is negative. What is needed for this possibility is that the sum be not *too small*—the negativity of $\mu(p, r; q) + \mu(p, r; \neg q)$ may so to speak not overpower the positivity of $\tau(p, r; q)$. So requirement (ii) is fulfilled if $\mu(p, r; q) + \mu(p, r; \neg q) > -\tau(p, r; q)$.

To show that we are talking about a real possibility see Fig. 1, which displays a probability distribution for which $0 > \mu(p, r; q) + \mu(p, r; \neg q) > -\tau(p, r; q)$.

From this probability distribution we calculate

$$\mu(p, r; q) = -\frac{2}{15} \quad \text{and} \quad \mu(p, r; \neg q) = -\frac{1}{50};$$

but nevertheless all the following differences are positive:

$$\begin{aligned} P(q|p) - P(q) &= \frac{17}{40} \\ P(r|q) - P(r) &= \frac{17}{48} \\ P(r|p) - P(r) &= \frac{7}{80}. \end{aligned}$$

How to define the condition of vw-screening off, which encompasses probability distributions like that of Fig. 1? It may seem that we have already found a suitable definition when we observed that $\mu(p, r; q) + \mu(p, r; \neg q)$ may be negative as long as this negativity does not swamp the positivity of $\tau(p, r; q)$. This line of thought yields as a candidate for the definition of vw-screening off:

$$\mu(p, r; q) + \mu(p, r; \neg q) > -\tau(p, r; q). \tag{11}$$

This inequality satisfies (i) and (ii): it is weaker than w-screening off, since it allows possibilities that are excluded by the latter, and it implies that p supports r . So it seems that (11) can take the rôle of C in (3):

$$(1) \rightarrow \left\{ (11) \rightarrow (2) \right\},$$

where (1) means that p supports q and q supports r , and (2) that p supports r .

However (11) is not acceptable, for it satisfies (3) only trivially: (11) by itself entails (2), there is no need for (1). To see this, add $\tau(p, r; q)$ to both sides of (11). Then we obtain

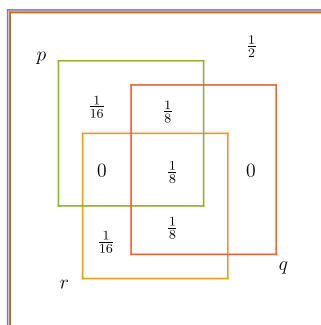


Fig. 1 Distribution in which $0 > \mu(p, r; q) + \mu(p, r; \neg q) > -\tau(p, r; q)$

$$\mu(p, r;q) + \mu(p, r;\neg q) + \tau(p, r;q) > 0.$$

From the identity (6) it then follows that $P(r|p) - P(r) > 0$ and thus that $P(r|p) > P(r)$. All that (11) does is to assert the tautology: $[P(r|p) > P(r)] \rightarrow [P(r|p) > P(r)]$.⁷

The fact that (11) does not need (1) in order to entail (2) goes against the very idea of transitivity, which after all is that p supports r through the mediator q . Clearly we require a condition for which (1) is needed.⁸ Note that both normal screening off and w-screening off satisfy this requirement: both need (1) in order to entail (2), since the conclusion that p supports r does not follow from (7) or (10) alone.

Here is a way to formulate a nontrivial condition of vw-screening off. Consider

$$\mu(p, r;q) + \mu(p, r;\neg q) + \tau(p, r;q) \geq \varepsilon\tau(p, r;q). \tag{12}$$

If $\varepsilon = 0$, then either (12) reduces to the trivial (11) or p does not support r at all. However, if $0 < \varepsilon < 1$, then (12) does the job. It then satisfies requirement (i), since $\mu(p, r;q) + \mu(p, r;\neg q)$ may be negative; it is thus weaker than w-screening off, which would correspond to (12) with $\varepsilon \geq 1$. It also satisfies (ii), for the Shogenji identity (6) shows that (12) is equivalent to $P(r|p) - P(r) \geq \varepsilon\tau(p, r;q)$; and because ε and $\tau(p, r;q)$ are both positive, p supports r . Finally, (12) with $0 < \varepsilon < 1$ is not a trivial condition. Since $\tau(p, r;q)$ could in general be negative, it needs (1) to ensure the positivity of $\tau(p, r;q)$:

$$\forall \varepsilon \in (0, 1) : (1) \rightarrow \{(12) \rightarrow (2)\} \quad \text{but} \quad \{(12) \nrightarrow (2)\}.$$

This is our condition of vw-screening off. Note that, because (12) contains ε explicitly, our condition for vw-screening off covers a continuum of conditions, one for each value of ε in the open interval (0, 1). Each of these conditions serves as a sufficient constraint for the transitivity of probabilistic support which is weaker than w-screening off. By making ε smaller and smaller, we can make the constraint as weak as we like.

In order to render our argument more intuitive and less abstract, we will in Sect. 5 give a real medical example of vw-screening off as defined by (12). But first, in

⁷ The same triviality lurks in a condition WC that William Roche introduces on p. 456 of Roche (2018). Transposed to our notation, WC is the statement “It is not the case that $\mu(p, r;q) + \mu(p, r;\neg q) < 0$ and $|\mu(p, r;q) + \mu(p, r;\neg q)| \geq \tau(p, r;q)$ ”. This is equivalent to the following disjunction

$$\left\{ \mu(p, r;q) + \mu(p, r;\neg q) \geq 0 \right\} \quad \text{or} \\ \left\{ \mu(p, r;q) + \mu(p, r;\neg q) < 0 \quad \text{and} \quad \mu(p, r;q) + \mu(p, r;\neg q) > -\tau(p, r;q) \right\}.$$

The first disjunct is the inequality (10), that is our condition of w-screening off; but the second disjunct is our trivial condition (11).

⁸ More precisely: what is needed is $\tau(p, r;q) > 0$, which is also consistent with $P(q|p) - P(q) < 0$ and $P(r|q) - P(r) < 0$. However, in this case $P(\neg q|p) - P(\neg q) > 0$ and $P(r|\neg q) - P(r) > 0$, i.e. p supports $\neg q$ and $\neg q$ supports r . We may say that $\tau(p, r;q) > 0$ is equivalent to (1), with an interchange of q and $\neg q$ where necessary (see footnote 12).

Sect. 4, we explain that vw-screening off has a remarkable property: it preserves transitivity even in the presence of the Simpson effect.

4 Transitivity and the Simpson Effect

The example of vw-screening given in Fig. 1 is also an instance of the Simpson effect. Recall the probability distribution in Fig. 1:

$$\mu(p, r; q) = -\frac{2}{15} \quad \text{and} \quad \mu(p, r; \neg q) = -\frac{1}{50} \quad \text{and} \quad P(r|p) - P(r) = \frac{7}{80}.$$

Here there is not only transitivity (p supports q , q supports r and p supports r), but there is also a Simpson effect: p *disconfirms* r conditionally on q and also conditionally on $\neg q$, but nevertheless p *confirms* r unconditionally.⁹

The fact that, via vw-screening off, transitivity can coexist with the Simpson effect is somewhat surprising—it was at least to us. For the two results seem to pull in different directions. To put it somewhat impressionistically, while transitivity has a ring of uninterruptedness to it, suggesting the continuous flow of probabilistic support through a chain, the Simpson effect gives the idea of an unexpected rupture, which is precisely why it is experienced as a paradox.

In general, the relation between the Simpson effect on the one hand and screening off (normal, weak, or very weak) on the other might appear somewhat complicated. Figure 2 can help us to achieve a better understanding of how precisely the two are connected. On the left, the smallest circle represents normal screening off (n-so) where $\mu(p, r; q) = 0$ and $\mu(p, r; \neg q) = 0$, the next circle represents weak screening off (w-so) where $\mu(p, r; q) + \mu(p, r; \neg q) \geq 0$, and the large circle represents very weak screening off (vw-so) where $\mu(p, r; q) + \mu(p, r; \neg q) > -\tau(p, r; q)$. In all of these three regions p supports q and q supports r , so $\tau(p, r; q)$ is positive and there is transitivity of probabilistic support.

The circle on the right represents the Simpson effect (se), in which also p supports r , but $\mu(p, r; q) < 0$ and $\mu(p, r; \neg q) < 0$. In the overlap region between vw-so and se, $\tau(p, r; q)$ is positive but $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ are both negative, and there is transitivity of support. Outside se and w-so but inside vw-so, one of $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ is positive and the other is negative, so there is no Simpson effect, but their sum is negative. But inside se and outside vw-so, $\mu(p, r; q)$ and $\mu(p, r; \neg q)$ are both negative and, although $\tau(p, r; q)$ is positive, p *disconfirms* q and q *disconfirms* r .

These properties have been implicitly obtained already in the literature. For example, the theorem in Appendix 1 of Lindley and Novick,¹⁰ translated into our notation, states that, ‘If Simpson’s paradox holds, with p and r positively correlated, and p and q positively correlated, then q and r are positively correlated’. This

⁹ Simpson (1951); Malinas and Bigelow (2016); Sprenger and Weinberger (forthcoming). In the latter, the Simpson effect is called ‘the association reversal’. Good and Mittal (1987) prefer ‘amalgamation paradox’. We will continue to speak of ‘the Simpson effect’ or ‘the Simpson reversal’.

¹⁰ Lindley and Novick (1981, 58).

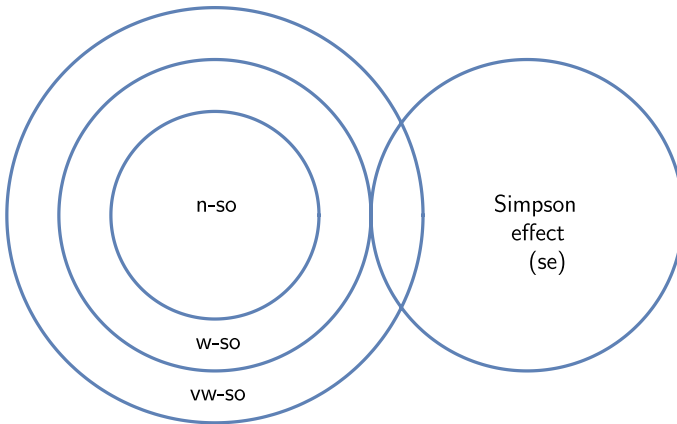


Fig. 2 Screening off and the Simpson effect

is one half of our finding. Mittal supplied what is equivalent to the other half of our result. Theorem 4.1 in Mittal’s paper, again translated into our notation, states that, ‘If Simpson’s paradox holds, with p and r positively correlated, then either (a) q is positively correlated to p and to r , or (b) q is positively correlated to $\neg p$ and to $\neg r$ ’.¹¹ Mittal’s alternative (b) amounts to $P(q|\neg p) > P(q)$ and $P(q|\neg r) > P(q)$, which is equivalent to $P(q|p) < P(q)$ and $P(r|q) < P(r)$. Thus Mittal’s case (b) corresponds to the region of Fig. 2 inside se but outside $vw-so$.

However, our approach enables us to see that there is more to be said. As we have seen, the Simpson effect does not imply that p supports q , and q supports r . Nevertheless it could be argued that the Simpson effect manifests a generalized sense of the transitivity of probabilistic support. For the Shogenji identity (6) can be transposed as follows:

$$\tau(p, r; q) = \left\{ P(r|p) - P(r) \right\} \left\{ -\mu(p, r; q) \right\} \left\{ -\mu(p, r; \neg q) \right\}.$$

Under Simpson reversal all the quantities between the parentheses {...} are positive, so $\tau(p, r; q)$ must be positive too. Thus either $P(q|p) > P(q)$ and $P(r|q) > P(r)$ or $P(q|p) < P(q)$ and $P(r|q) < P(r)$. However in the latter case $P(\neg q|p) > P(\neg q)$ and $P(r|\neg q) > P(r)$.¹² So p supports $\neg q$, and $\neg q$ supports r . So whenever Simpson reversal occurs, p supports r , either through the mediation of q or through the mediation of $\neg q$.

Various estimates of probabilistic support, under the guise of Bayesian measures of the confirmation of hypotheses, have been listed by Fitelson (1999) and others. Although they suffer from the disadvantage that they are not ordinally equivalent to one another, they do all agree that, if $P(p \wedge r) > P(p)P(r)$, then $c(r, p) > 0$, and if $P(p \wedge r) < P(p)P(r)$, then $c(r, p) < 0$, where $c(r, p)$ is any of the aforementioned

¹¹ Mittal (1991, 171). Mittal calls the Simpson paradox ‘Yule’s Reversal Paradox’, giving credit to Yule (1903), see Mittal (1991, 168).

¹² Firstly $[P(q|p) < P(q)] \longrightarrow [P(\neg q|p) = 1 - P(q|p) > 1 - P(q) = P(\neg q)]$; and secondly $[P(r|q) < P(r)] \longrightarrow [P(q|r) < P(q)] \longrightarrow [P(\neg q|r) > P(\neg q)] \longrightarrow [P(r|\neg q) > P(r)]$.

Bayesian measures of confirmation. Accordingly, whenever the Simpson effect occurs, then one or other of the following alternatives is true:

- (a) $c(r, p) > 0$ and $c(q, p) > 0$ and $c(r, q) > 0$
- (b) $c(r, p) > 0$ and $c(q, p) < 0$ and $c(r, q) < 0$.

5 Example: Kidney Stones

A real life instance of the Simpson effect concerns the removal of kidney stones (renal calculi). Julious and Mullee drew attention to a study that had been made by Charig and coworkers of the success rates of two kinds of operations to remove the stones: open surgery or percutaneous nephrolithotomy (the penetration of the skin and kidney by a tube, through which the stone is removed).¹³

Julious and Mullee concentrated on 700 operations that were performed on patients with kidney stones, one half by open surgery (between 1972 and 1980), and the other half by percutaneous nephrolithotomy (between 1980 and 1985). An operation was deemed successful if no stones greater than 2 mm in diameter were present in the operated kidney three months after the operation; and success rates were compared for stones that were smaller or larger than 2 cm in diameter.

Consider one operation among these 700, and define the following propositions:

- r : the operation was successful
- p : percutaneous nephrolithotomy was performed
- $\neg p$: open surgery was performed
- q : the stone that was removed was less than 2 cm in diameter
- $\neg q$: the stone that was removed was at least 2 cm in diameter

Since the number of percutaneous nephrolithotomies was equal to the number of open surgeries (namely 350), $P(p) = 0.5$.

The numbers given by Charig et al. correspond to the following conditional probabilities (relative frequencies):¹⁴

$$\begin{aligned}
 P(r|p) &= 0.83 & P(r|\neg p) &= 0.78 \\
 P(r|p \wedge q) &= 0.87 & P(r|\neg p \wedge q) &= 0.93 \\
 P(r|p \wedge \neg q) &= 0.69 & P(r|\neg p \wedge \neg q) &= 0.73
 \end{aligned}
 \tag{13}$$

The complete probability distribution can be extracted from these numbers: it has been reproduced in Fig. 3. From this distribution we calculate

$$P(r|p) - P(r) = 0.025,$$

¹³ Julious and Mullee (1994) and Charig et al. (1986).

¹⁴ Julious and Mullee incorrectly give the percentage of successes for percutaneous nephrolithotomies with stones of diameter less than 2 cm as 83%, whereas according to Charig et al. it should be 87%. This is presumably a copying error, for with 83% the probability distribution would be inconsistent whereas with 87% the distribution is consistent and there is indeed a Simpson effect.

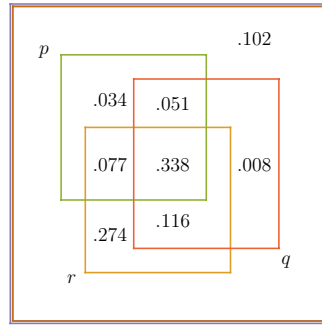


Fig. 3 Distribution for the kidney stone operations

so p supports r , i.e. percutaneous nephrolithotomy improves the chance of success. On the other hand,

$$\begin{aligned} P(r|q \wedge p) - P(r|q) &= -0.016 \\ P(r|\neg q \wedge p) - P(r|\neg q) &= -0.003. \end{aligned} \tag{14}$$

So percutaneous nephrolithotomy *decreases* the chance of success for stones of less than 2 cm diameter, and *also* for stones at least as large as 2 cm. This is of course an example of the Simpson effect, which was the burden of the paper of Julious and Mullee.

From Fig. 3 we can also calculate

$$\begin{aligned} P(q|p) - P(q) &= 0.265 \\ P(r|q) - P(r) &= 0.080, \end{aligned}$$

thus p supports q , and q supports r , in the sense that the correlations in question are positive. In other words, the kidney stone example displays the Simpson effect and also transitivity of probabilistic support; it is in fact an instance of very weak screening off.

Eqs. (14) imply that the two μ functions are negative:

$$\mu(p, r; q) = -0.0125 \quad \mu(p, r; \neg q) = -0.0060,$$

while the τ function is positive:

$$\tau(p, r; q) = 0.0435.$$

The sum

$$\mu(p, r; q) + \mu(p, r; \neg q) + \tau(p, r; q) = 0.025$$

is equal to $P(r|p) - P(r)$, as should be the case.

We see from (6) and (12) that

$$\epsilon \leq \frac{P(r|p) - P(r)}{\tau(p, r; q)} = \frac{0.025}{0.0435} = 0.57.$$

6 A Necessary Condition

We started our inquiry by recalling the well-known fact that probabilistic support is in general not transitive. Tomoji Shogenji however showed that it *is* transitive under (normal) screening off, and William Roche proved that normal screening off can be weakened, while retaining the transitivity; earlier proofs can be found in Eells and Sober (1983), and Suppes (1986). In this paper we offered an alternative proof of their results, making use of a powerful identity in Shogenji (2017).

We then weakened weak screening off further to what we called very weak screening off, defined by inequality (12) where ϵ satisfies $0 < \epsilon < 1$. Very weak screening off covers a continuum of conditions, each of which is weaker than weak screening off and is sufficient for transitivity. We pointed out that a special case of very weak screening off includes a special case of the Simpson effect, which we illustrated by means of an example about kidney stones taken from the seminal paper by Julious and Mullee.

Like normal and weak screening off, very weak screening off is a nontrivial sufficient condition for (2), given (1). If $0 < \epsilon < 1$, then

$$(1) \rightarrow \{(12) \rightarrow (2)\} \quad \text{but} \quad \{(12) \nrightarrow (2)\}.$$

However, by placing an extra constraint on ϵ we can turn (12) into a nontrivial necessary condition as well. With any ϵ satisfying the inequality

$$0 < \epsilon \leq 1 + \frac{\mu(p, r; q) + \mu(p, r; \neg q)}{\tau(p, r; q)}, \tag{15}$$

we will show that:

$$(1) \rightarrow \{(12) \leftarrow (2)\} \quad \text{but} \quad \{(12) \nleftarrow (2)\}.$$

According to the Shogenji identity, the right-hand side of (15) is equal to

$$\frac{\mu(p, r; q) + \mu(p, r; \neg q) + \tau(p, r; q)}{\tau(p, r; q)} = \frac{P(r|p) - P(r)}{\tau(p, r; q)}. \tag{16}$$

From (1), $\tau(p, r; q)$ is positive, and from (2), $P(r|p) - P(r)$ is positive, so it follows that the right-hand side of (16) is positive, so the left-hand side is positive, too. Note that (2) is required to make the right-hand side of (15) positive: without (2) ϵ could be negative, which would make (15) inconsistent. On multiplying (15) throughout by $\tau(p, r; q)$, we obtain

$$\epsilon \tau(p, r; q) \leq \mu(p, r; q) + \mu(p, r; \neg q) + \tau(p, r; q),$$

and this is none other than (12).

We can now draw the following three conclusions.

1. For any $\epsilon > 0$, inequality (12) is a sufficient and nontrivial condition for transitivity. The Shogenji identity (6) tells us that (12) is equivalent to

$$P(r|p) - P(r) \geq \epsilon \tau(p, r; q),$$

and since $\tau(p, r; q)$ could be negative, (12) is nontrivial in the sense that it does not imply (2) by itself: (12) $\not\leftrightarrow$ (2). However, if (1), then $\tau(p, r; q)$ is positive. With ϵ also being positive, (12) suffices for transitivity: (1) $\rightarrow \{(12) \rightarrow (2)\}$.

2. The domain $\epsilon > 0$ can be divided into two subdomains:
 - (a) $\epsilon \geq 1$, when weak screening off applies (or normal screening off as a special case);
 - (b) the more stringent $\epsilon < 1$, when very weak screening off holds sway. Very weak screening off encompasses weak and normal screening off, but also includes probability distributions that fall outside the domain of weak screening off.

In either case (a) or (b) transitivity transpires, which is why we were able simply to specify $\epsilon > 0$ as the constraint for sufficiency.

3. The even stronger restriction

$$0 < \epsilon \leq 1 + \frac{\mu(p, r; q) + \mu(p, r; \neg q)}{\tau(p, r; q)}$$

is a nontrivial necessary and sufficient condition for transitivity, that is for the following to hold,

$$(1) \rightarrow \{(12) \leftrightarrow (2)\} \quad \text{but} \quad \{(12) \leftrightarrow (2)\}.$$

Note that, since the inequality (12) depends on ϵ , there is in effect a separate condition of necessity and sufficiency for each value of ϵ in the permitted range. In examining various conditions for transitivity, we have been relying on the \forall quantifier: we talked about particular values for all ϵ in a particular domain. As an alternative, one could employ the quantifier \exists . For example, one could express the necessary and sufficient condition for probabilistic support as:

$$\exists \epsilon > 0 : (1) \rightarrow \{(12) \leftrightarrow (2)\} \quad \text{but} \quad \{(12) \leftrightarrow (2)\}.$$

This expression has the advantage of being economical, indeed terse.¹⁵ The disadvantage might be that it hides much under the logical carpet.

Acknowledgements We are very grateful for the discerning criticism and advice of three anonymous reviewers. Thanks also to William Roche and Tomoji Shogenji for earlier discussions on transitivity.

¹⁵ We thank an anonymous reviewer for making this suggestion.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Carnap, R. (1962). *Logical foundations of probability* (2nd ed.). Chicago: University of Chicago Press.
- Charig, C. R., Webb, D. R., Payne, S. R., & Wickham, J. E. A. (1986). Comparison of treatment of renal calculi by open surgery, percutaneous nephrolithotomy, and extracorporeal shockwave lithotripsy. *British Medical Journal*, *292*, 879–882.
- Crupi, V., Tentori, K., & Gonzalez, M. (2007). On Bayesian measures of evidential support: Theoretical and empirical issues. *Philosophy of Science*, *74*, 229–259.
- Eells, E., & Sober, E. (1983). Probabilistic causality and the question of transitivity. *Philosophy of Science*, *50*, 35–57.
- Fitelson, B. (1999). The plurality of Bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science*, *66*, 363–378.
- Good, I. J., & Mittal, Y. (1987). The amalgamation and geometry of two-by-two contingency tables. *The Annals of Statistics*, *15*, 694–711.
- Julious, S., & Mullee, M. (1994). Confounding and Simpson's paradox. *British Medical Journal*, *309*, 1480–1481.
- Lindley, D. V., & Novick, M. R. (1981). The role of exchangeability in interference. *The Annals of Statistics*, *9*, 45–58.
- Malinas, G., & Bigelow, J. (2016). Simpson's paradox. The Stanford encyclopedia of philosophy (Winter 2016 Edition), E. N. Zalta (Ed.). <https://plato.stanford.edu/archives/fall2016/entries/paradox-simpson/>.
- Mittal, Y. (1991). Homogeneity of subpopulations and Simpson's paradox. *The Journal of the American Statistical Association*, *86*, 167–172.
- Roche, W. (2012). A weaker condition for transitivity in probabilistic support. *European Journal for the Philosophy of Science*, *2*, 111–118.
- Roche, W. (2018). Is evidence of evidence evidence? Screening-Off vs. No-Defeaters. *Episteme*, *15*, 451–462.
- Shogenji, T. (2003). A condition for transitivity in probabilistic support. *The British Journal for the Philosophy of Science*, *54*, 613–616.
- Shogenji, T. (2017). Mediated confirmation. *The British Journal for the Philosophy of Science*, *68*, 847–874.
- Simpson, E. (1951). The interpretation of interaction in contingency tables. *Journal of the Royal Statistical Society (Series B)*, *13*, 238–241.
- Sprenger, J., & Weinberger, N. (forthcoming). Simpson's paradox. *The Stanford Encyclopedia of Philosophy*.
- Suppes, P. (1986). Non-Markovian causality in the social sciences with some theorems on transitivity. *Synthese*, *68*, 129–140.
- Yule, G. U. (1903). Notes on the theory of association of attributes in statistics. *Biometrika*, *2*, 121–134.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.